



PROSIDING
SEMINAR NASIONAL
KE 7 Tahun 2012
Rekayasa Teknologi Industri dan Informasi
Inovasi Teknologi dan Informasi untuk
Optimalisasi Energi

SEKOLAH TINGGI TEKNOLOGI NASIONAL
YOGYAKARTA

SUSUNAN PANITIA

Penanggung Jawab	: Ketua STTNAS
Pengarah	: Pembantu Ketua
KetuaPelaksana	: Ir. Harianto, MT.
Sekretaris Pelaksana Staff Sekretariat	: Ir. Eka Yawara, MT. : 1. Sri Harjanti : 2. Sunah
BendaharaPelaksana	: Drs. Sukapdi
SeksiMakalah	:
Koordinator	: Dr. Hill. Gendoet Hartono, ST., MT.
Teknik Mesin	: Dr. Ratna Kartikasari, ST, MT.
Teknik Elektro	: Tugino, ST, MT.
Teknik Sipil	: Drs. H. Triwuryanto, MT.
Teknik Geologi	: Dr. Ir. Ev. Budiadi, MS.
Teknik PWK	: Drs. Achmad Wismoro, ST, MT.
Teknik Pertambangan	: Ir. Ag. Isjudarto, MT.
Seksi Proseeding	: 1. Ir. Muhammad Abdulkadir, MT. : 2. Djoko Purwanto, ST.
Seksi Acara	: Sigit Budi Hartono, ST, MT.
Seksi Publikasi, Dokumentasi	: 1. ArisWarsita, ST, MT. : 2. Ferry Okto Satriya, ST. : 3. Ign. Purwanto : 4. H. Andiyanto, Amd.
Sponsor	: 1. Ir. Nizam Effendi : 2. Sulaiman Tampubolon, ST.

DAFTAR ISI

SUSUNAN PANITIA	ii
SAMBUTAN KETUA PANITIA ReTII KE 7	iii
SAMBUTAN KETUA STTNAS	iv
DAFTAR ISI	v

TEKNIK ELEKTRO

1. Penggunaan Algoritma Differential Evolution Dalam Penyelesaian Kombinasi Pembebanan Optimal Ekonomis Dan Emisi Pada Pembangkit Listrik Termal <i>Afner Saut Sinaga</i>	1
2. Kendali Level Kecepatan Motor DC Lima Tingkat dengan Rheostat (Resistance Control) Terintegrasi Safety Deadman Pedal Pada Sistem Kereta Api Berbasis PLC (Programmable Logic Control) <i>Arifin Wibisono, Jefri Setiawan, Leonardus Heru Pratomo</i>	7
3. Pengaruh Trafik Paket Aplikasi terhadap Kinerja Jaringan dengan Manajemen Bandwidth Fifo pada Warnet Rush Yogyakarta <i>Ayu Budi Setyawati, Damar Widjaja</i>	11
4. Pengembangan Indoor Location Based Service Menggunakan Wireless Positioning pada Android <i>Dwijayanto Gusti Parrangan, Y. Sigit Purnomo Wuryo Putro, B. Yudi Dwiandiyanta</i>	17
5. Power Monitoring Berbasis Mikrokontroler <i>Freddy Kurniawan</i>	23
6. Sistem Pemerolehan Informasi Makalah Ilmiah Berbahasa Indonesia Menggunakan Struktur Data Inverted Index Berbasis Ordbms Dengan Metode Pembobotan Tf-Idf <i>Justina S. Wulandari, JB Budi Darmawan</i>	29
7. Kendali Buck-Boost Mppt Berbasis Digital <i>Matias Chosta Agryatma, Slamet Riyadi, F. Budi Setiawan</i>	35
8. Sistem Penjejak Lokasi Sumber Suara Menggunakan Interaural Time Difference <i>Muhammad Afridon, Djoko Purwanto</i>	39
9. Sistem Pemerolehan Informasi Dokumen Makalah Ilmiah Berbahasa Indonesia Menggunakan Struktur Data Inverted Index Berbasis Hash Table Dan Ordered Linkedlist <i>Reza M. Darojad, JB Budi Darmawan</i>	45
10. Desain Kontroler Fuzzy Logic untuk Robot Pembersih Sampah dalam Ruang <i>Tri Hendrawan Budianto, Irwan Dinata</i>	51
11. Kombinasi Vb dan Matlab untuk Pemrosesan Sinyal Radar Ransponder Rocket <i>Wahyu Widada</i>	57
12. Optimasi Kerja Baterai Charge-Discharge pada Sistem Pengaturan Beban (Power Management) di BTS (Base Transceiver Station) Remote Area Menggunakan Pengaturan Beban Dinamis <i>Widjonarko</i>	61
13. Perancangan Konverter Energi Berbasis Buck Chopper Untuk Panel Surya <i>Y. L. Christanto Wibowo, Ign Slamet Riyadi</i>	69
14. Desain Sistem Pembangkit Listrik Tenaga Surya Beroperasi Stand Alone dengan Konverter Ky dan Maximum Power Point Tracking Berbasis Algoritma Neuro-Fuzzy <i>Adi Kurniawan, Mochamad Ashari, Dedet C. Riawan, Ilham Pakaya</i>	75
15. Rancang Bangun Water-Meter Digital dengan Transfer Data Melalui Short Message Service (SMS) <i>Joko Prasajo, Arif Basuki, Armansyah</i>	81
16. Peningkatan Kualitas Citra Digital Dengan Metode Non-Linear Filter <i>Agus Basukesti</i>	87
17. Estimasi Kanal MIMO OFDM Berdasarkan Perubahan Nilai Signal to Noise Ratio (SNR) <i>Anggun Fitriani Isnawati</i>	93

18. Pengembangan Robot Pengikut Garis Berbasis Logika Fuzzy <i>Aji Joko Budi Pramono</i>	101
19. Perancangan Boost Konverter Sebagai Interface Antara Panel Surya Dan Beban <i>Fx Anton Yk Slamet Riyadi</i>	107
20. Pengaruh Berbagai Ekstraksi Ciri Terhadap Tingkat Pengenalan Isyarat Tutar pada Sistem Pengenal Tutar Model Markov Tersembunyi <i>Asniar Aliyu</i>	113
21. Brushless Direct Current (BLDC) Motor Controller Using Digital Logic For Electric Vehicle <i>Bambang Sujanarko</i>	121
22. Desain dan Implementasi Maksimal Power Point Tracker dengan Kendali Tenganan untuk Sistem Pengisi Baterai <i>Banar Arianto , Leonardus. H. Pratomo</i>	125
23. Analisa Ekonomi Pemasangan Distributed Generation PLTU Prafi II Manokwari <i>Elias K. Bawan, Pandung Sarungallo</i>	131
24. Strategi Untuk Membantu Eksekutif Dalam Pengambilan Keputusan Dengan Menggunakan Data Warehouse Pengadaan Pelumas Pada PT. ABC <i>Evaristus Didik. M, Dewi. S, Felisia. L, Winnie. S</i>	137
25. Analisa Penggunaan Home Solar Cell untuk Efisiensi Pemakaian Energi Listrik (Implementasi Pemakaian Pada Perumahan Type 27/66) <i>Irfan Santoso, Tofik Hidayat</i>	143
26. Pengenalan Nada Pianika Menggunakan Fft Dan Korelasi <i>Dionysius Edwin Surya, Linggo Sumarno</i>	151
27. Maximum Power Point Tracking Menggunakan Artificial Neural Network Untuk Sistem PV Terhubung Grid Melalui Wide Range Input Inverter <i>Muhammad Syafei Gozali, Dedet Candra Riawan, Mochamad Ashari</i>	159
28. Rancang Bangun Alat Bantu Penentuan Lokasi Kincir Angin Pada PLT ANGIN <i>Tito Yuwono, Budi Astuti, Febrian Fariz</i>	165
29. Pengembangan E-Procurement dengan Menggunakan Kerangka Kerja Cobit (Studi Kasus : Kementerian Keuangan Timor - Leste) <i>Onorio Dos Santos, Benyamin L. Sinaga, Paulus Mudjihartono</i>	171
30. Perancangan Catu Daya Dengan High Frequency Transformator Berbasis Kendali Digital <i>Dionisius Wahyu Pradana, Ign. Slamet Riyadi</i>	177
31. Dummy Load Untuk Beban 450 Watt <i>Pernandes, Martanto</i>	181
32. Penentuan Tegangan Penyalaan (Ignition-Voltage) pada Kendaraan Tangki Pembawa Bahan Bakar Cair Premium Dan Gas Lpg dalam Kawasan Bermedan Listrik <i>Budi Utama</i>	187
33. Desain Sistem Jaring Kecil (Micro Grid System) Berbasis Photovoltaic Menggunakan Kontrol Artificial Neural Network <i>Wan Muhammad Faizal, M. Ashari, Heri Suryo.A</i>	195
34. Chopper-Inverter Sebagai Interface Pv dan Sistem Kelistrikan <i>Ricky Gondo Atmodjo, Slamet Riyadi,</i>	201
35. Perbandingan Kinerja Empat Metode Prototipe Alat Ukur Kadar Curcuminoid pada Rimpang Kunyit (Curcuma Domestica) <i>Bernadeta Wuri Harini, Rini Dwiastuti, Lucia Wiwid Wijayanti</i>	205
36. Analisis Pengaruh Penempatan UPFC Terhadap Minimisasi Rugi Saluran Transmisi <i>Petrus Setyo Prabowo</i>	211
37. Analisa dan Perancangan Portal Web Untuk Konsultan Pajak <i>Stefanus Cendra Hogi Sopacua, Flourensia Sapty Rahayu, Eduard Rusdianto</i>	217
38. Prototipe Multigain Gyroscope Untuk Aplikasi Roket <i>Sri Kliwati</i>	223
39. Rancang Bangun Robot Animaloid Berkaki Empat	

<i>Tugino, Septian Andra, Sudiana</i>	227
40. Pengaruh Tegangan Impuls terhadap Ketahanan Arester Tegangan Rendah <i>Diah Suwarti</i>	231
41. Perancangan Sistem Embedded berbasis FPGA <i>Totok Mujiono, Tasripan, Pujiono</i>	237
42. Desain Sistem Dual Inputs Sepic – Bidirectional Converter untuk Manajemen Energi Sistem Pembangkit Photovoltaic Pada Area Terpencil <i>Daniar Fahmi, Dedet C. Riawan, M. Ashari</i>	241
43. Kelayakan Penerapan Sistem Kendali Suhu Dan Kelembaban Pada Budidaya Jarum Tiram di Jogjakarta <i>Hendra Setiawan, Sholichin</i>	249

TEKNIK MESIN

1. Model Peningkatan Kualitas Layanan Kesehatan di Jawa Timur Melalui Integrasi Metode Servqual, Lean dan Six Sigma Untuk Meningkatkan Kepuasan Konsumen <i>Hana Catur Wahyuni, Wiwik Sulistiyowati</i>	255
2. Pengaruh Penambahan Krom dan Tempo terhadap Kekerasan, Keausan dan Struktur Mikro Ball Mill <i>Sumpena, Subarmono, R. Soekrisno</i>	262
3. Perancangan Dan Pembuatan Mesin Produksi Palet Ikan Dengan Pendekatan Ergonomis <i>Tofik Hidayat, Irfan Santoso</i>	269
4. Pengaruh Kadar Si Terhadap Struktur Mikro Dan Sifat Mekanik Kandidat Baja Ringan Paduan Fe-Al-Mn <i>Ratna Kartikasari, Sutrisna</i>	275
5. Studi Peningkatan Daya dan Torsi dengan Pemasangan Air Tube pada Saluran Intake Manifold Sepeda Motor Dua Langkah <i>Harjono</i>	280
6. Pengembangan Program Pendukung Keputusan Untuk Estimasi Manufacturing Cost Pada Perancangan Cold Storage Menggunakan Panel Surya <i>Boni Sena, Fauzun, Endang Suhendar</i>	284
7. Aplikasi Interferometer Michelson Untuk Pengukuran Regangan Pada Mesin Uji Tari <i>Budi Setyahandana, Martanto, Ronny Dwi Agusulistyo</i>	289
8. Efek Perubahan Ukuran Diameter Header Knalpot terhadap Konsumsi Bahan Bakar dan Akselerasi Kendaraan Pada Motor 4 Tak <i>Aji Pranoto</i>	296
9. Penentuan Jalur Terpendek Petugas Kebersihan Sampah Di Lingkungan Perumahan Dosen UGM Sekip Menggunakan Algoritma Semut <i>Andhi Akhmad Ismail, Radhian Krisnaputra</i>	302
10. Pengaruh Perubahan Debit Aliran Udara-Air Terhadap Respon Amplitudo dengan Menggunakan Gelombang Ultrasonik Pada Aliran Stratified Horizontal <i>Mukhlis, Bramantya, Hermawan</i>	308
11. Pengaruh Volume Ruang Bakar Terhadap Kinerja Mesin Pulse Jet <i>Lambertus Dwi Setiawan</i>	314
12. Studi Eksperimen Batas Mampu Bakar Campuran LPG / CO ₂ sebagai Refrigeran Alternatif <i>Nasrul Iminnafik</i>	318
13. Menyelidiki Pengaruh Pemasangan Pelat-Pelat Pengarah Angin pada Keliling Lingkaran Luar Sudu Kincir Terhadap Unjuk Kerja Tiga Model Kincir Angin Savonius <i>Rines</i>	322
14. Identifikasi, Pemodelan dan Kompensasi Ketidaktelitian Open Loop Control System Pada Mesin Milling CNC Mini <i>Ignatius Aris Hendaryanto, M. Arif Wibisono, Herianto</i>	329
15. Analisis Pengerasan Permukaan Metode Flame Hardening dengan Pencekaman Spesimen Sistem Vertikal Pada Baja S45C	

SISTEM PEMEROLEHAN INFORMASI DOKUMEN MAKALAH ILMIAH BERBAHASA INDONESIA MENGGUNAKAN STRUKTUR DATA INVERTED INDEX BERBASIS HASH TABLE DAN ORDERED LINKEDLIST

Reza M. Darojad¹, JB Budi Darmawan²

^{1,2}Jurusan Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Sanata Dharma
Kampus III Paingan Maguwohardjo Depok Sleman Yogyakarta

¹virtual.024@gmail.com , ²jbbudi@gmail.com

Abstrak

Makalah ilmiah yang dipublikasikan di Indonesia bertambah banyak, sehingga kalangan publik terutama akademik memerlukan suatu sistem pemerolehan informasi yang dapat membantu mereka untuk menemukan informasi yang relevan. Pencarian dokumen makalah ilmiah yang sudah ada sebagian besar menghasilkan informasi berdasarkan dari abstrak, judul, pengarang, penerbit, dan subjek makalah. Dalam penelitian ini penulis mengimplementasikan Sistem Pemerolehan Informasi menerapkan pengindekan menggunakan seluruh isi teks makalah ilmiah berbahasa Indonesia. Sistem ini menggunakan metode *TF-IDF* dan struktur data *classical* yaitu *Inverted Index* yang berbasis pada *Hash Table* dan *Ordered LinkedList*. Hasil ujicoba menggunakan 281 dokumen menunjukkan unjuk kerja waktu akses yang diperoleh untuk 1 kata kunci 0,001 detik hingga 0,033 detik dan untuk 4 kata kunci 0,001 detik hingga 0,0491 detik.

Kata Kunci : Pemerolehan Informasi, *Inverted Index*, *Hash*, *LinkedList*

1. PENDAHULUAN

1.1 Latar belakang

Jumlah dokumen makalah ilmiah digital di Indonesia terus bertambah sehingga masyarakat terutama akademisi memerlukan suatu sistem yang dapat mengakses dan menyediakan berbagai informasi sesuai kebutuhannya. Informasi tersebut dapat diperoleh menggunakan sistem temu-kembali informasi (*Information retrieval*) agar pengguna mendapatkan sumber informasi yang tepat sesuai kebutuhan pengguna dari sekumpulan dokumen yang besar [1].

TF-IDF adalah salah satu metode dari *Information Retrieval* untuk memberikan bobot hubungan suatu kata (*term*) terhadap dokumen. Metode *TF-IDF* yang digunakan menggunakan teknik pembobotan *Savoy* karena pada teknik pembobotan ini, bobot istilah telah dinormalisasi. Teknik ini memperhitungkan jumlah dokumen yang mengandung istilah yang bersangkutan dan jumlah keseluruhan dokumen. Sehingga jika sebuah istilah mempunyai frekuensi kemunculan yang sama pada dua dokumen belum tentu mempunyai bobot yang sama. Teknik ini sudah diterapkan pada dokumen berbahasa Indonesia dan cukup baik dalam memberikan bobot dokumen terurut[6].

Data yang berisi daftar term dan hubungannya dengan dokumen disimpan dalam sebuah *Inverted Index*. *Inverted index* terdiri dari dua bagian utama, *dictionary* dan *posting list*. Struktur data *Hash Table* adalah pilihan sesuai digunakan sebagai *dictionary* karena waktu aksesnya relatif singkat dimana pengurutan data tidak diutamakan. *Posting*

List memerlukan struktur data penyimpanan yang terurut dan dinamis, *Ordered Linked List* diimplementasikan untuk *Posting List* karena ukuran *Linked List* yang dapat melebar menyesuaikan data yang ditambahkan. *Ordered Linked List* berdasarkan pada *Linked List* dengan penambahan pengurutan data[3].

Pencarian dokumen makalah ilmiah yang sudah ada sebagian besar menghasilkan informasi berdasarkan abstrak, judul, pengarang, penerbit, dan subjek makalah. bukan seluruh isi makalah. Peneliti mengembangkan Sistem Pemerolehan Informasi untuk dokumen makalah berbahasa Indonesia menggunakan seluruh isi teks dokumen.

1.2 TUJUAN

Penelitian ini bertujuan untuk mengimplementasikan sistem pemerolehan informasi untuk pencarian makalah ilmiah berbahasa Indonesia menggunakan seluruh isi teks makalah ilmiah dan mengetahui unjuk kerja sistem yang menggunakan *Inverted Index* klasik dengan struktur data *Hash Table* dan *Ordered LinkedList* dengan operasi *boolean* dasar.

1.3 METODE PENELITIAN

Dalam penelitian ini dilakukan tahap-tahap penelitian sebagai berikut:

1. Studi pustaka penerapan *inverted index* klasik menggunakan pembobotan *TF-IDF* dengan operasi AND.

2. Pengumpulan dokumen-dokumen makalah ilmiah yang berbahasa Indonesia sebagai *corpus*.
3. Implementasi penerapan *TF-IDF* dan struktur data *classical* yaitu *Inverted Index* yang berbasis pada *Hash* dan *Ordered LinkedList* untuk mendukung *inverted index* dengan menggunakan rumus pembobotan Savoy.
4. Pengamatan unjuk kerja waktu *query* dengan operasi AND dengan dua belas kelompok kata yang mengandung frekuensi dokumen tertentu. Kelompok kata tersebut adalah kelompok kata yang memiliki dfk 1 sampai 2, dfk mendekati 140, dan dfk kurang lebih 280, 1 kata kunci pencarian hingga 4 kata kunci pencarian.

Penelitian ini dilakukan dengan menggunakan komputer yang memiliki spesifikasi sebagai berikut:

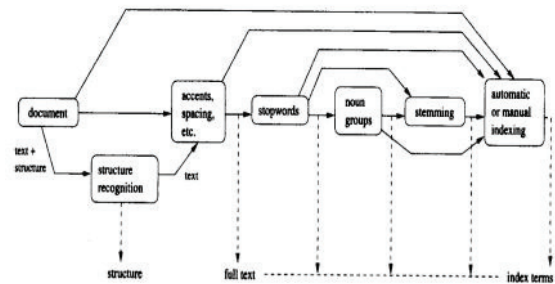
- a. Perangkat Lunak
 1. Sistem Operasi Oracle Enterprise Linux 5.8
 2. Oracle 11g R2 Enterprise Edition
 3. Oracle SQL Developer 3.0.04
 4. Java JDK 1.6.0
 5. Netbeans 6.9.1
- b. Perangkat Keras
 1. Prosesor Xeon E5620(4 Core, 2.4 Ghz)
 2. RAM 8 GB RDIMM
 3. Hardisk RAID 5 logikal 2 TB
 4. Motherboard HP Proliant ML 350 G6

2. LANDASAN TEORI

2.1 Information Retrieval

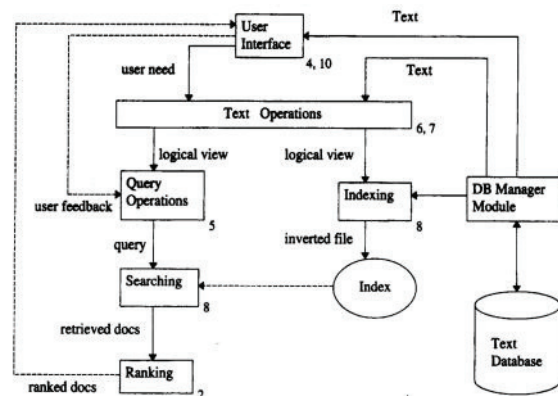
Sistem temu kembali informasi (*information retrieval system*) merupakan sistem untuk menemukan kembali (*retrieve*) informasi-informasi yang relevan terhadap kebutuhan pengguna dari suatu kumpulan informasi secara otomatis. Penekanannya ada pada penemukembalian informasi yang sifatnya tidak terstruktur. Salah satu contoh dari sistem temu kembali informasi adalah *search-engine* atau mesin pencarian [5].

Dokumen diwakili melalui *set index term*. *Index term* menyediakan *logical view* dari dokumen. Jika koleksi dokumen cukup besar komputer akan melakukan pengurangan jumlah *set term* melalui penghapusan *stopwords*, operasi *stemming*. Operasi teks tersebut akan mengurangi kompleksitas dari dokumen dan mengubah *logical view* dari *full text* ke *set index term*. Gambar 1 menunjukkan *logical view* yang digunakan sistem pemerolehan informasi[6].



Gambar 1. Logikal view dari sebuah dokumen: dari *full text* menjadi sebuah set indeks *term*[6].

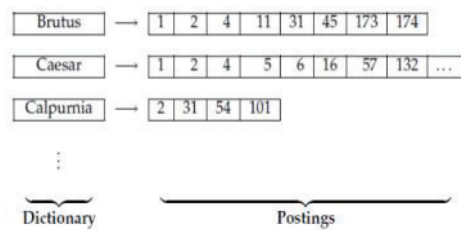
Sistem pemerolehan informasi memiliki beberapa tahap. Pertama melakukan indentifikasi terhadap dokumen-dokumen yang akan digunakan, operasi yang akan dilakukan terhadap teks, dan model teks. *Text operations* mentransformasikan dokumen asal menjadi *logical view*. Setelah *logical view* diperoleh, dibuat indeks *term* untuk mempercepat pencarian terhadap jumlah data yang besar. Setelah dokumen selesai diindeks, proses *retrieval* dapat diinisiasi. Sebelum dikirim ke pengguna, dokumen *retrieval* di peringkat terlebih dahulu berdasarkan kemungkinan relevansinya.[7]



Gambar 2. Proses dari pemerolehan informasi[6]

2.2 Inverted Index

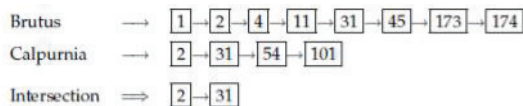
Inverted index terdiri dari dua bagian utama, *dictionary* dan *posting list*. *Dictionary* menyimpan daftar kata, sedangkan *posting list* menyimpan identitas dokumen yang mengandung kata yang bersangkutan. Setiap kata terhubung dengan satu rangkaian *posting list* yang terurut menggunakan penunjuk[1].



Gambar 3. Dua bagian dari *Inverted Index*[1].

Untuk memproses *query* AND pada *inverted index* sebagai contoh “Brutus AND Calpurnia”, pada Gambar 3 berikut langkahnya:

1. Temukan lokasi Brutus dalam *Dictionary*.
2. *Retrieve posting* dari Brutus.
3. Temukan lokasi Calpurnia dalam *Dictionary*.
4. *Retrieve posting* dari Calpurnia.
5. Ambil hanya dokumen yang terdapat di kedua *posting list* tersebut dengan melakukan interseksi, seperti pada gambar 4[1].



Gambar 4. Interseksi *posting lists*[1].

2.3 Pembobotan TF-IDF

Teknik pembobotan Savoy (1993) sebagai berikut[5]

$$W_{ik} = \text{ntf}_{ik} * \text{nidf}_k,$$

$$\text{dimana } \text{ntf}_{ik} = \frac{tf_{ik}}{\text{Max}_j tf_{ij}} \text{ dan } \text{nidf}_k = \frac{\log\left[\frac{n}{df_k}\right]}{\log(n)}$$

$$Wd = Wij + Wik + Wih \dots + Win$$

- W_{ik} adalah bobot istilah k pada dokumen i .
- tf_{ik} merupakan frekuensi dari istilah k dalam dokumen i .
- n adalah jumlah dokumen dalam kumpulan dokumen.
- df_k adalah jumlah dokumen yang mengandung istilah k .
- $\text{Max}_j tf_{ij}$ adalah frekuensi istilah terbesar pada satu dokumen.
- Wd = bobot sebuah dokumen

Pada teknik pembobotan ini, bobot istilah telah dinormalisasi. Dalam menentukan bobot suatu istilah tidak hanya berdasarkan frekuensi kemunculan istilah di satu dokumen, tetapi juga memperhatikan frekuensi terbesar pada suatu istilah yang dimiliki oleh dokumen bersangkutan. Hal ini untuk menentukan posisi relatif bobot dari istilah dibanding dengan istilah-istilah lain di dokumen yang sama. Selain itu teknik ini juga memperhitungkan jumlah dokumen yang mengandung istilah yang bersangkutan dan jumlah keseluruhan dokumen. Hal ini berguna untuk mengetahui posisi

relatif bobot istilah bersangkutan pada suatu dokumen dibandingkan dengan dokumen-dokumen lain yang memiliki istilah yang sama. Sehingga jika sebuah istilah mempunyai frekuensi kemunculan yang sama pada dua dokumen belum tentu mempunyai bobot yang sama.[5]

2.4 Struktur Data

Hash Table adalah struktur data yang menawarkan pemasukan dan pencarian data dengan sangat cepat. Ide dari *Hash Table* adalah memperbolehkan banyak dari kemungkinan *key* berbeda yang mungkin di petakan ke lokasi yang sama di dalam *array* dibawah fungsi tindakan pengindeksan. *Hash Table* beroperasi relatif cepat $O(1)$ kali. Untuk pembuatan daftar kamus kata, *hash table* sesuai untuk digunakan[3].

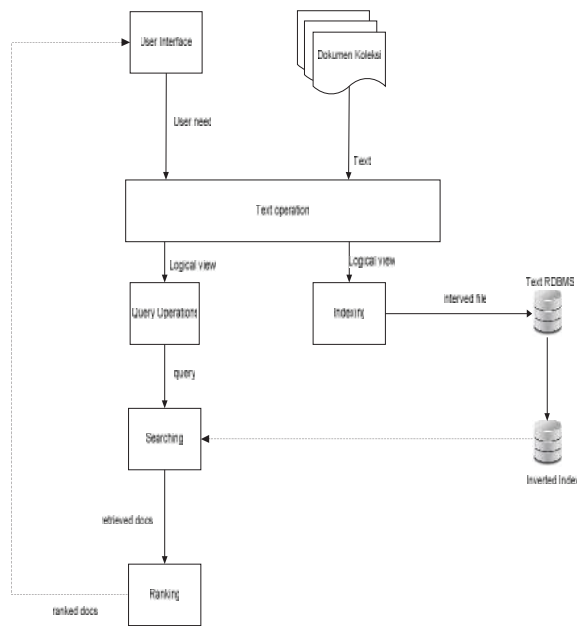
Linked List adalah struktur data yang berbentuk *node* yang menunjuk *node* lainnya menggunakan *link*[2]. Ukuran *Linked List* menjadi dinamis karena ukurannya bertambah mengikuti jumlah *node* yang dimasukkan kedalam rantai *node*[3].

3. Pembahasan dan Hasil Penelitian

3.1 Arsitektur Sistem

Arsitektur sistem pemerolehan informasi ditunjukkan pada Gambar 5. Dokumen koleksi berupa file makalah ilmiah berbahasa Indonesia dengan ekstensi .pdf. Sistem melakukan *text operation* pada dokumen tersebut, berupa pembuangan kata buang dan *stemming* menggunakan algoritma Nazief & Adriani untuk teks berbahasa Indonesia, lalu melakukan proses *indexing* kata dan dokumen pada RDBMS. Sistem mengambil *index* dari RDBMS ke random access *memory* secara periodik di dalam *Inverted Index* serta melakukan pembobotan.

Pada proses pencarian pengguna memasukkan *query*, yang akan diproses oleh *text operation*. Sistem melakukan pencarian pada *Inverted Index* dengan operasi AND kemudian melakukan *ranking* berdasarkan bobot dokumen. Hasil dokumen yang sudah memiliki peringkat ditampilkan melalui *user interface*.

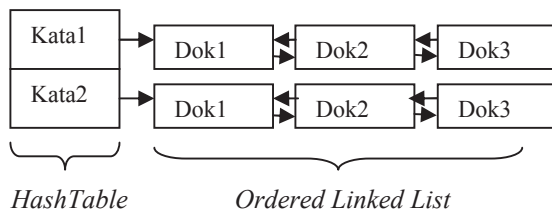


Gambar 5. Rancangan Sistem Pemerolehan Informasi.

3.2 Data yang Digunakan

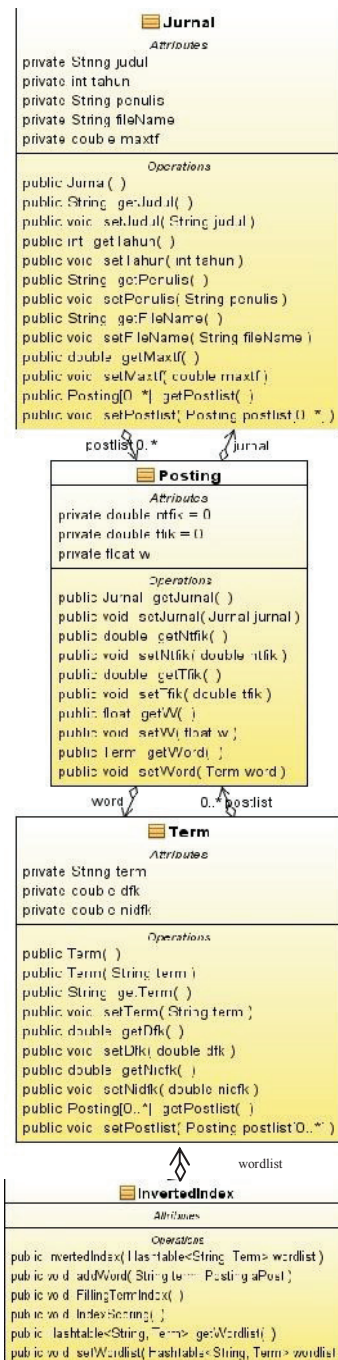
Inverted Index yang dibuat menggabungkan antara *Hash Table* sebagai *dictionary* dan *Ordered Linked List* sebagai *posting list* seperti pada Gambar 6.

Ordered Linked List berdasarkan pada *Linked List*, namun dilakukan *sorting* pada saat akan menambahkan data sehingga *posting* dimasukkan secara terurut berdasarkan *id document*.



Gambar 6. Representasi *Inverted Index*

Untuk *Inverted Index* dimodelkan dalam diagram kelas pada Gambar 7. Pada kelas *Term* data yang disimpan adalah *term* yaitu kata, *dfk* untuk frekuensi dokumen, dan *nidf* untuk *inverse dokumen frequency* ternormalisasi. Pada kelas *Posting*, data yang disimpan adalah *ntfk* untuk *term frequency* ternormalisasi, *tf* untuk frekuensi kata, dan *w* untuk bobot. Pada kelas *Jurnal* data yang disimpan adalah judul, tahun, penulis, nama *file*, dan *maxtf* untuk frekuensi kata maksimum dalam satu dokumen.



Gambar 7. Rancangan model *Inverted Index*.

3.3 Hasil Percobaan

Percobaan menggunakan operasi AND dengan melakukan interseksi antar *postinglist* yang diperoleh saat pencarian ke dalam *Inverted Index*.

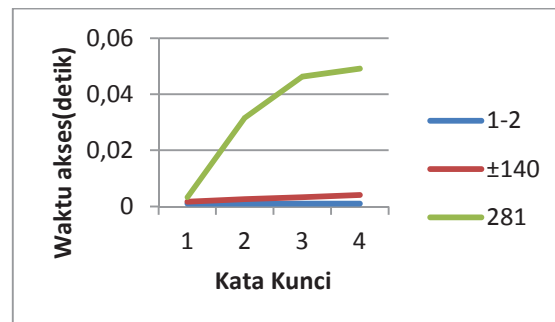
Pada Tabel 1 menunjukkan pada kelompok *dfk* 1-2 untuk untuk kelompok 1 hingga 4 kata kunci berada pada 0,001 detik, hal ini karena sistem masih belum bisa mencatat waktu kurang dari 0,001 detik, kemungkinan waktu tersebut kurang dari 0,001 detik. Pada kelompok *dfk* ± 140 , waktu akses diperoleh 0,0017 detik untuk 1 kata kunci

hingga 0,0041 detik untuk 4 kata kunci. Pada grafik Gambar 8 terlihat peningkatan waktu akses tersebut *linear*. Pada kelompok dfk 281 terjadi peningkatan waktu akses dari 0,0033 detik untuk 1 kata kunci hingga 0,0491 detik untuk 4 kata kunci. Peningkatan waktu akses dari 1 kata kunci ke 2 kata kunci cukup signifikan dibandingkan dari 2 sampai 4 kata kunci, hal ini disebabkan ada operasi AND untuk 2,3, dan 4 kata kunci sedangkan pada kelompok 1 kata kunci tidak ada.

Dari Gambar 8 dapat dilihat pada kelompok dfk ± 140 peningkatan waktu akses *linear* untuk peningkatan dari 2 sampai 4 kata kunci. Pada kelompok dfk 281 peningkatan waktu akses pada 2 kata kunci ke 3 kata kunci cukup tajam dibandingkan peningkatan waktu akses dari 3 kata kunci ke 4 kata kunci, hal ini disebabkan jumlah dokumen relevan yang diperoleh untuk 3 kata kunci sama dengan 4 kata kunci yaitu 281. Sedangkan pada kelompok dfk ± 140 terjadi penurunan hasil dokumen yang relevan dari 63 dokumen untuk 3 kata kunci menjadi 41 untuk 4 kata kunci.

Tabel 1. Waktu rata-rata *query* untuk kelompok kata 1-2 dfk, ± 140 dfk, dan 281 dfk dengan operasi AND.

DFK	Jumlah Kata Kunci	Waktu Query Rata2 (detik)	Jumlah Dokumen yang Relevan
1-2	1	0,001	1
	2	0,001	2
	3	0,001	1
	4	0,001	1
± 140	1	0,0017	138
	2	0,0026	85
	3	0,0033	63
	4	0,0041	41
281	1	0,0033	281
	2	0,0316	281
	3	0,0463	281
	4	0,0491	281



Gambar 8. Grafik Waktu rata-rata *query* untuk kelompok kata 1-2 dfk, ± 140 dfk, dan 281 dfk operasi AND.

Untuk kata kunci 1 sampai 4, kelompok dfk 281 memiliki waktu akses paling tinggi, diikuti kelompok dfk ± 140 , dan yang terakhir kelompok dfk 1-2. Hal ini menunjukkan peningkatan jumlah dokumen untuk setiap *postinglist*(dfk) meningkatkan waktu akses *query*. Waktu akses untuk 1 kata kunci 0,001 detik hingga 0,033 detik dan untuk 4 kata kunci 0,001 detik hingga 0,0491 detik.

4. Penutup

4.1 Kesimpulan

Sistem pemerolehan informasi yang menggunakan Struktur data *Inverted Index* berbasis *Hash Table* dan *Ordered Linked List* ini memiliki unjuk kerja pencarian 0,001 detik untuk 1 hingga 4 kata kunci pada kelompok kata dfk 1-2. Unjuk kerja sistem 0,0017 detik untuk 1 kata kunci hingga 0,0041 detik pada kelompok kata dfk ± 140 dengan grafik peningkatan waktu akses *linear*. Unjuk kerja sistem 0,0033 detik untuk 1 kata kunci hingga 0,0491 detik untuk 4 kata kunci pada kelompok kata maksimum 281.

Waktu akses yang diperoleh untuk 1 kata kunci 0,001 detik hingga 0,033 detik dan untuk 4 kata kunci 0,001 detik hingga 0,0491 detik. Peningkatan frekuensi dokumen untuk setiap kata kunci ikut mempengaruhi waktu yang diperlukan untuk proses pencarian yang juga meningkat.

4.2 Saran

Penelitian selanjutnya dapat dikembangkan untuk operasi pencarian OR dan NOT atau operasi kombinasi AND, OR, NOT.

DAFTAR PUSTAKA

- [1] Manning, Christopher D. (2009). *An Introduction to Information Retrieval*. Cambridge University Press Cambridge, England
- [2] Kruse, Robert L. (1994). *Data Structure and Program Design Third Edition*. Prentice Hall International, Inc, London.

- [3] Robert, Lafore (2003). *Data Structure & Algorithms in Java Second Edition*. Sams Publishing, Indiana.
- [4] A.H , Muh. Erwin & Mandala, Rila. (2004), *Relevance Feedback pada Temu Kembali Informasi Menggunakan Algoritma Genetika*. Seminar Nasional Aplikasi Teknologi Informasi 2004, Yogyakarta.
- [5] Hasibuan, Zainal A., & Andri, Yofi. (2001). *Penerapan Berbagai Teknik Sistem Temu-Kembali Informasi Berbasis Hiperteks*. Jurnal Ilmu Komputer dan Teknologi Informasi, Volume 1, Nomor 2.
- [6] Baeza-Yates, R., Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. Addison Wesley.